Supplement of

Singularity-sensitive gauge-based radar rainfall adjustment methods for urban hydrological applications

L.-P. Wang et al.

Correspondence to: L.-P. Wang (lipen.wang@bwk.kuleuven.be)

The copyright of individual parts of the supplement might differ from the CC-BY 3.0 licence.
FORMULATION AND DEMONSTRATION OF THE SINGULARITY-SENSITIVE BAYESIAN DATA MERGING METHOD

The singularity-sensitive method was initially developed and preliminarily tested in the reconstruction of a storm event which led to reported pluvial flooding of a number of properties in the Maida Vale area of London (part of the Counters Creek catchment) on the 27th June 2009 between 15:00 and 20:15. The radar rainfall product - UK Met Office Nimrod composite (Golding, 1998; Harrison et al., 2000) - for this event showed strong and localised singularity structures; however, when inputted into a calibrated hydraulic model of the area, no flooding was generated. Given that the hydraulic model had been shown to be reliable, the problem seemed to lie with the accuracy of the radar rainfall estimates. To resolve the accuracy problem while preserving the spatial structure of the rainfall field as captured by the radar, a dynamic gauge-based adjustment was conducted using the Bayesian data merging method, which in previous studies had been shown to outperform other adjustment methods (Mazzetti and Todini, 2004; Wang et al., 2013). Nonetheless, for this particular event, records from only few rain gauge sites were available and these were located away from the area of interest and at points where less intense radar rainfall was observed, leaving critical areas (where higher intensities and accumulations were recorded by the radar) uncovered. Figure 1 shows the locations of the rain gauges over a 28 x 28 km² area covering the Counters Creek catchment, as well as the cumulative rainfall depths for 15:00-20:15 period on 2009/06/27 from the Nimrod radar data set. Under these circumstances, the aforementioned shortcomings associated with existing adjustment techniques became evident. The Bayesian data merging method proved inadequate as it smoothed out the singular structures, which had the effect of considerably reducing the peak rainfall intensities. In view of this, the proposed method of incorporating a singularity analysis into the original Bayesian merging procedure was developed.

In what follows, a description is first provided of the two key techniques used in this paper: the Bayesian data merging method and the local singularity analysis. Afterwards, the proposed method for integrating these two techniques is explained. In this supplementary document, the radar and rain gauge data of the 2009 Maida Vale event are used to illustrate intermediate results and highlight the special features of the techniques under investigation. Due to confidentiality issues, the hydraulic model for this catchment and the associated hydraulic outputs could not be presented in the paper.
Bayesian radar-rain gauge data merging method

The Bayesian data merging method (BAY) is a dynamic adjustment method (applied independently at each time step) intended for real-time applications (Todini, 2001). The underlying idea of this method is to analyse and quantify the uncertainty of rainfall estimates (in terms of error co-variance model) from multiple data sources – in this case, radar and rain gauge sensors – and then combine these estimates in such a way that the overall (estimation) uncertainty is minimised. The BAY merging method consists of the following steps (illustrated in Figure 2(a)):

a. For each time step $t$, the point rain gauge (RG) measurements are interpolated into a synthetic rainfall field using the block-kriging (BK) technique. The result of this step is an interpolated rain gauge rainfall field, with areal estimates at each radar grid location ($y_{t}^{RG}$) and which are accompanied by the associated estimation error co-variance function ($V_{t}^{ε_{RG}}$), representing the uncertainty of rain gauge estimates.

b. The interpolated rain gauge rainfall field is compared against the radar field ($y_{t}^{RD}$), based upon which a field of errors (estimated as the bias at each radar grid location: $ε_{t}^{RD} = y_{t}^{RD} - y_{t}^{RG}$) is obtained empirically. Assuming that areal rain gauge estimates are unbiased, the expectation value ($μ_{t}^{ε_{RD}} = E[ε_{t}^{RD}]$) and the co-variance function ($V_{t}^{ε_{RD}}$) of this radar-rain gauge error field at each time step is used to represent, respectively, the mean bias and the uncertainty of radar estimates.

c. Using a Kalman filter (Kalman, 1960), the two rainfall fields are optimally combined such that the overall estimation uncertainty is minimised. In the Kalman filter the radar data and the interpolated rain gauge estimates act, respectively, as ‘a priori estimate’ and ‘measurement’. The degree of ‘uncertainty’ of each type of estimate constitutes a gain value (the so-called Kalman gain, $K_{t}$) at each radar grid location, and determines the proportion of each type of estimate that is used to compute the merged output. The use of this gain value ensures the
minimisation of the overall estimation uncertainty and is expressed as
\[ K_t = V_{\varepsilon t}^{RD}(V_{\varepsilon t}^{RD} + V_{\varepsilon t}^{RG})^{-1}. \] (1)

and the optimally BAY merged output (i.e. the ‘a posteriori’ estimates \( y'_t'^{''} \) in the Kalman filter) can be obtained from
\[ y'_t'^{''} = y'_t' + K_t(y_t^{RG} - y'_t'). \] (2)

where \( y'_t' \) is the ‘unbiased’ radar rainfall estimate (i.e. \( y'_t' = y_t^{RD} - \mu_{\varepsilon t}^{RD} \)), used as the ‘a priori’ estimate in the Kalman filter.

It can be seen that the Kalman gain is a function of the error co-variances of radar and rain gauge estimates. When \( V_{\varepsilon t}^{RG} \gg V_{\varepsilon t}^{RD} \) (or \( K_t \approx 1 \), i.e. radar estimates have significantly higher uncertainty than the rain gauge ones), the radar estimates are trusted less and the output estimates will be very similar to the interpolated rain gauge field. In contrast, when \( V_{\varepsilon t}^{RG} \gg V_{\varepsilon t}^{RD} \) (or \( K_t \approx 0 \)), the output will be closer to the radar estimates.

It is in steps b and c where the problems associated with the Bayesian merging technique, and geostatistical techniques in general, arise. The (2\textsuperscript{nd})-order co-variance function that these techniques employ to characterise radar-rain gauge errors cannot capture local singularity structures. Instead, in 2\textsuperscript{nd}-order models singularities are mistakenly regarded as errors in the radar data, thus leading to higher estimated radar uncertainty, \( V_{\varepsilon t}^{RD} \). In the BAY method higher \( V_{\varepsilon t}^{RD} \) causes the radar data to be trusted less; this ultimately results in smoother merged outputs, which are closer to the interpolated rain gauge field.

Radar (RD), BK and BAY rainfall rate estimates over the Counters Creek catchment at the time of peak intensity (2009/06/07 16:55) are shown in Figure 4 (a1), (b1) and (c1). Noteworthy is the smoothness of the BAY estimates, and its similarity with the BK field, which has the effect of considerably reducing the peak rainfall intensities initially observed in the RD image. Particularly critical is the centre-left area of the domain (i.e. around the area of interest – see Figure 1), where intense rainfall cells were observed in RD, yet there was no rain gauge coverage in this area. The lack of rain gauge measurements in this area led to BK and associated BAY estimates showing nearly zero rainfall, thus losing the important spatial features of the rainfall field initially captured by the radar. This highlights the sensitivity of the BK and BAY methods to the number and location of rain gauges. Moreover, it exposes the effect of the underlying Gaussian approximation, which causes the BAY technique to give more weight to the ‘smooth’ and ‘reliable’ estimates generated by the BK technique and subsequently neglect the local peaks in the RD data, which may be regarded as significant errors. The smoothing carried out by the BK interpolation and BAY merging method over the entire domain is further reflected in making the field approximately normally distributed, with smaller probabilities of extremes (particularly large ones), as compared to the original RD field. This can be seen in Figure 4 (a2), (b2) and (c2), where the quantiles of the observed distribution against the quantiles of a normal distribution with the same mean and variance are shown.
Local singularity analysis

Various types of hazardous geo-processes, including precipitation, often result in anomalous amounts of energy release or mass accumulation confined to narrow intervals in time and/or space. The property of anomalous amounts of energy release or mass accumulation is termed a singularity and it is often associated to structures depicting fractality or multifractality (Agterberg, 2007; Cheng, 1999; Lovejoy and Mandelbrot, 1985; Schertzer and Lovejoy, 1987). Several mathematical models and methodologies have been developed to respectively characterise and treat singularities. In this work, the local singularity analysis proposed by (Cheng et al., 1994) has been adopted to identify and extract singularities from rainfall fields. Cheng’s method, which has been widely used for estimation of geo-chemical concentrations (Agterberg, 2007; Cheng and Zhao, 2011; Cheng et al., 1994), employs the definition of coarse Hölder exponent to characterise singularities. According to this model, singularities are defined through the fact that the areal average measure (in this work, areal rainfall) centred at a specific location $x$ (taken as the centre of a radar pixel) varies as a power function of the area $E$. This power-law relationship can be formulated as an equation (Cheng et al., 1994):

$$\rho(x, \epsilon) = c(x)e^{\alpha(x) - E},$$

(3)

where $\rho(x, \epsilon)$ represents the density of measure (e.g. concentration of geo-data; in the context of this paper, rainfall intensity) over a square area with side-length $l$ and associated scale $\epsilon$ ($\epsilon = l/L$, where $L$ is the side-length of the largest square area under consideration) centred at a specific location $x$; $c(x)$ is a constant value (in the context of this paper, a constant intensity value) at $x$; $\alpha(x)$ is the singularity index/exponent (or the coarse Hölder exponent); and $E=2$ is the Euclidean dimension of the plane.

A schematic of the estimation of the constant value $c(x)$ and singularity index $\alpha(x)$ from real gridded data is provided in Figure 3. For a given pixel with centre $x$ (centre of top plot in Figure 3), the ‘mean’ rainfall intensities at different spatial scales (centred in $x$) can be calculated (i.e. rainfall intensities $\rho_1$, $\rho_2$, ..., $\rho_n$, respectively at scales $\epsilon_1$, $\epsilon_2$, ..., $\epsilon_n$). Then, the logarithms of these mean values and the associated spatial scales are compared (bottom plot in Figure 3). The constant value $c(x)$ and singularity index $\alpha(x)$ of the dataset can be derived by applying a simple linear regression analysis, where the ‘slope’ and the ‘y-intercept’ of the regression line correspond, respectively, to the terms $(\alpha(x) - E)$ and $\log c(x)$. A detailed explanation of the computation of $c(x)$ and $\alpha(x)$ can be found in previous studies (Agterberg, 2012; Chen et al., 2007; Cheng et al., 1994). It is worth mentioning that the estimation of $c(x)$ and $\alpha(x)$ can be trusted only if a good
linear relation is observed (i.e. if the scaling behaviour is well followed).

Going back to the definition of singularity, Eq (3) constitutes a useful tool to decompose an areal rainfall intensity at a given location \( x \) into two components (Wang et al., 2012): (1) the background (or non-singular, NS) magnitude \( c(x) \), which is invariant as measuring scale \( \epsilon \) changes and is more approximately normal than the original field, and (2) a local ‘scaling’ multiplier, the magnitude of which changes as the measuring scale \( \epsilon \) changes, according to the local singularity index \( a(x) \). When \( a(x) < 2 \), the rainfall magnitude strikingly increases as the measuring scale \( \epsilon \) decreases (namely local enrichment); this corresponds to a ‘peak’ singularity. In contrast, when \( a(x) > 2 \), the rainfall magnitude decreases as \( \epsilon \) decreases (i.e. local depletion), and it is therefore a ‘trough’ singularity. When \( a(x) = 2 \), there is no singularity: the rainfall intensity \( \rho(x, \epsilon) \) within a \( \epsilon \times \epsilon \) area remains the same as scale changes (i.e. \( \rho(x, \epsilon) = c(x) \)).

There is, however, a drawback to this local singularity analysis in practice. Because it carries out a ‘local’ analysis, the singularity exponents are usually obtained from a small number of data samples. This increases the uncertainty of the estimation of \( a(x) \). The consequence of this drawback is that the singularity is incorrectly estimated or incompletely extracted; therefore, \( c(x) \) is an unreliable or incomplete non-singular value. To circumvent this, two numerical strategies were employed in this paper. The first one involves constraining the value of the estimated singularity exponents within a certain range. This can avoid obtaining unreasonably large or small singularity exponents. A number of ranges, symmetric to the non-singular condition (i.e. \( a(x) = 2 \)), were selected for testing. They are (from the widest to narrowest intervals): \( \text{SIN1} = [0, 4] \), \( \text{SIN2} = [0.5, 3.5] \), \( \text{SIN3} = [1, 3] \), \( \text{SIN4} = [1.5, 2.5] \) and \( \text{SIN5} = [1.75, 2.25] \). These ‘truncated’ singularity ranges were empirically chosen according to the authors’ experience and the fact that the distribution of \( a(x) \) is seldom largely skewed to a specific side of \( a(x) = 2 \) (see Figure 5 (b)). It can be generally expected that, the wider the range is, the more singularity information (both local enrichment and depletion) from the radar images is taken into account in the merging process.

The second numerical strategy is to decompose the rainfall field using an iterative procedure (Agterberg, 2012; Chen et al., 2007):
\[ c^{(k-1)}(x) = c^{(k)}(x)e^{\alpha^{(k)}(x)-E}, \]

Where the iterative index \( k = 0, 1, 2, \ldots, n \). When \( k = 0 \), \( c^{(-1)}(x) = \rho(x, \epsilon) \) (i.e. the original value) and \( c^{(0)}(x) \) is the ‘calculated’ ‘non-singular’ value from the first iteration, which is equal to \( c(x) \) from the non-iterative calculation above (Eq. (4)). This \( c^{(0)}(x) \) is then used as the left-hand-side value of Eq. (4) to calculate the ‘non-singular’ value at the next iteration, and so on. Substituting Eq. (4) into Eq. (3), one can obtain an iterative local singularity analysis equation:

\[ \rho(x, \epsilon) = c^*(x)e^{\alpha^*(x)-E}, \]

where

\[
\begin{align*}
    c^*(x) &= c^{(n)}(x) \\
    \alpha^*(x) &= \alpha^{(0)}(x) + \sum_{k=1}^{n}(\alpha^{(k)}(x) - E).
\end{align*}
\]

The criterion to terminate the iteration procedure is when \( \alpha^{(k)}(x) \approx E \) (which is equivalent to \( c^{(k-1)}(x) \approx c^{(k)}(x) \)). That means the singularity components have been clearly removed from the data.

Moreover, in this work a spatial-scale range of 1-9 km, which results in a total of five rainfall intensity samples (at scales 1, 3, 5, 7 and 9 km), was used in the singularity analysis. This range was selected for two main reasons. Firstly, our analyses revealed that a good linear behaviour was generally observed within this scale range, while small-scale structures were still preserved in the resulting rainfall product. As such, the selected spatial scale range was deemed to represent a good balance between estimation uncertainty (which depends upon the number of samples employed in the calculations) and local feature preservation. Secondly, a scaling break at approximately 8-16 km has been reported in studies in which 1-km radar rainfall data were analysed (Gires et al., 2012; Tchiguirinskaia et al., 2011). This means that the rainfall data at spatial-scale regimes ranging from 1 to 16 km comply with the same or similar statistical or physical behaviour. This scaling range has also been used in other applications to represent relatively local characteristics of rainfall fields (Bowler et al., 2006).

Lastly, a 10-iteration singularity analysis was applied in order to ensure that most of the singularity exponents could be extracted. The trade-off of conducting many iterations is the longer computational time, which may be an issue for real-time applications. Nonetheless, in practice, approximately 4-6 iterations are sufficient for effectively removing most of the singularity.

An example of the application of the local singularity analysis (with a singularity range \( \text{SIN3} = [1, 3] \)) to the decomposition of the 2009/06/27 peak intensity radar image for the Counters Creek catchment (comprising the Maida Vale area) can be found in Figure 4(d). It can be seen that, as compared to the original radar (RD) image (Figure 4(a1)), the spatial structure of the non-singular component \( c(x) \) (Figure 4(d1): NS-RD) is smoother, more symmetric and spreading over a larger area. In addition, the NS-RD estimates are much closer to normality than the original RD data. In Figure 4(d2), it can be seen that the NS-RD estimate quantiles are highly consistent with the Normal theoretical quantiles. This is not the case for the original RD estimates, where a much longer tail at the right end of the data distribution is observed (Figure 4(a2)). Given its ‘Normal’ behaviour, the NS-RD estimates may be a more suitable input than the original RD estimates for many existing data merging techniques based upon the Gaussian approximation.

**Incorporation of the local singularity analysis into the Bayesian merging method**

The underlying idea of the proposed method is to use the local singularity analysis to decompose each radar image into a non-singular image \( c^*(x) \) and a singularity map \( \alpha^*(x) \) before applying the Bayesian merging (i.e. step (i) in Figure 2 (b)). The non-singular radar image (NS-RD), which has a distribution closer to normality (thus being more suitable for Gaussian-based treatments), is merged with the point rain gauge data following the Bayesian procedure. This yields a
non-singular Bayesian merged field (NS-BAY). Afterwards, the singularity map is applied back and proportionally to the NS-BAY merged field (i.e. step (ii) in Figure 2 (b)), thus yielding a singularity-sensitive merged field (SIN). This is done by multiplying each pixel value of the NS-BAY field by the following ratio:

\[
r(x) = \epsilon^{a(x) - E},
\]

which corresponds to the ratio difference between the original radar field (RD) and the non-singular radar field (NS-RD). In this way, a singularity-sensitive merged field (SIN), which better retains the local singularity structures embedded in the original RD field is obtained.

NS-BAY and SIN merged estimates over the Counters Creek catchment at the time of peak intensity (2009/06/07 16:55) are shown in Figure 4 (e1) and (f1). The associated quantile-quantile plots are shown in Figure 4 (c2) and (f2). By comparing the NS-BAY (Figure 4 (e1)) and the BAY (Figure 4 (c1)) merged estimates, it can be seen that the former retains more features of the RD image, particularly in the centre-left area of the domain, where no rain gauge measurements were available. This means that the reliability of the original RD estimates was effectively increased by removing the singularities from it; therefore, the NS-RD estimates were trusted more in the BAY merging process. Moreover, it can be seen that after the singularities are applied back to the NS-BAY, yielding the SIN estimates, the missing local rainfall information and the non-normality of the original RD rainfall field are effectively recovered (Figure 4 (f1) and (f2)). In fact, as compared to the BAY estimates, the SIN merged estimates show a less smooth and more realistic spatial structure. What is more, the SIN estimates successfully preserve the rainfall cell structure captured by the RD in the problematic and critical centre-left area of the domain, where, as mentioned above, no rain gauge measurements were available.

It is worth noting that the proposed singularity-sensitive merging method does not always increase the reliability of RD estimates. Such increase only happens when the RD estimates exhibit high singularity and thus cannot be well handled using Gaussian approximations. This feature of the proposed method can be illustrated by analysing the variation of RD estimation uncertainty before and after the singularity exponents are extracted from RD estimates at different time points of the storm event. Two RD images, at 16:55 (a1) and 18:00 (a2) (respectively inside and outside the peak period of the 2009/06/27 event), are compared in Figure 5. In the histogram of singularity exponents of each image (Figure 5(b)), it can be seen that nearly 70% of the singularity exponents in the 18:00 image are equal to 2, while only 45% of them in the 16:55 image are equal to 2. This means that the 16:55 image is of much higher singularity than the 18:00 one. In Figure 5(c1) and (c2), the BK-RD errors (the cross markers) correspond to the original RD error co-variance function (i.e. the \(V_e^{rd} \) term in Eq. (1)), while the BK-NS errors (the round markers) represent the NS-RD error co-variances (i.e. after the local singularity analysis has been applied to the original RD image (i.e. step (i) in Figure 2 (b)), yielding a non-singular RD image, NS-RD). It can be observed that the variation of these two co-variance functions is much more significant for the 16:55 image than for the 18:00 image. This indicates that the local singularity analysis is particularly effective in reducing the uncertainty of RD estimates when highly singular spatial structures are present.

Based upon the intermediate results shown in this section, it can be concluded that, as compared to the original BAY method, the proposed SIN method has better ability to handle local extreme magnitudes that cannot be well characterised nor preserved by most existing merging techniques. This advantage could be in particular critical for small urban-scale hydrological modelling.
Figure 4: Top: Snapshot images at the time of peak intensity (29/06/2009 16:55) over the Counters Creek catchment - (a1) radar estimates (RD), (b1) block-kriged rain gauge estimates (BK), (c1) original Bayesian merged estimates, (d1) non-singular radar estimates (NS-RD), (e1) non-singular Bayesian merged estimates, (f1) singularity-sensitive Bayesian merged estimates (SIN). Bottom: Associated quantile-quantile (Q-Q) plots (a2-f2) showing the quantiles of the observed distribution against the quantiles of a normal distribution with the same mean and variance (the closer the dots are to a straight line, the more normally distributed the data are).

Figure 5: (a1, a2) Radar (RD) images, respectively at 16:55 and 18:00 of the 2009/06/27 event; (b) associated histogram of singularity exponents ($\alpha$) for each of the two radar images; (c1, c2) radar error co-variance functions, respectively at 16:55 and 18:00, where BK-RD errors (cross markers) represent the co-variances of biases between BK and the (original) RD estimates, and the BK-NS errors (round markers) are those between BK and NS-RD (non-singular RD) estimates.

REFERENCES


Tchiguirinskaia, I., Schertzer, D., Hoang, C. T. and Lovejoy, S.: Multifractal study of three storms with different dynamics over the Paris region, in 12th International Conference on Urban Drainage, Porto Alegre, Brazil., 2011.

